Two-Sided Matching with Partial Information

BAHARAK RASTEGARI, University of British Columbia
ANNE CONDON, University of British Columbia
NICOLE IMMORLICA, Northwestern University
KEVIN LEYTON-BROWN, University of British Columbia

The traditional model of two-sided matching assumes that all agents fully know their own preferences. As markets grow large, however, it becomes impractical for agents to precisely assess their rankings over all agents on the other side of the market. We propose a novel model of two-sided matching in which agents are endowed with known partially ordered preferences and unknown true preferences drawn from known distributions consistent with the partial order. The true preferences are learned through interviews, revealing the pairwise rankings among all interviewed agents, performed according to a centralized interview policy, i.e., an algorithm that adaptively schedules interviews. Our goal is for the policy to guarantee both stability and optimality for a given side of the market, with respect to the underlying true preferences of the agents. As interviews are costly, we seek a policy that minimizes the number of interviews. We introduce three minimization objectives: (very weak) dominance, which minimizes the number of interviews for any underlying true preference profile; Pareto optimality, which guarantees that no other policy dominates the given policy; and optimality in expectation with respect to the preference distribution. We formulate our problem as a Markov decision process, implying an algorithm for computing an optimal-in-expectation policy in time polynomial in the number of possible preference orderings (and thus exponential in the size of the input). We then derive structural properties of dominant policies which we call *optimality certificates*. We show that computing a minimum optimality certificate is NP-hard, suggesting that optimal-in-expectation and/or Pareto optimal policies could be NP-hard to compute. Finally, we restrict attention to a setting in which agents on one side of the market have the same partially ordered preferences (but potentially distinct underlying true preferences), and in which agents must interview before matching. In this restricted setting, we show how to leverage the idea of minimum optimality certificates to design a computationally efficient interview-minimizing policy. This policy works without knowledge of the distributions and is dominant (and so is also Pareto optimal and optimal-in-expectation).

Categories and Subject Descriptors: J.4 [**Social and Behavioral Sciences**]: Economics

Additional Key Words and Phrases: Matching; Market design; Partial information

1. INTRODUCTION

Two-sided matching markets model many practical settings, such as corporate hiring, marriage, and university admission. In such markets, participants are partitioned into two disjoint sets—e.g., men and women in a marriage market. Each participant on one side of the market wishes to be matched to a candidate from the other side of the market, and has preferences over potential matches. A matching is called stable if no pair of participants would prefer to leave their assigned partners to pair with each other. A matching is optimal for one side of the market if each participant on the given side prefers the matching to any other stable matching. Gale and Shapley [1962] proposed a tractable algorithm for identifying optimal, stable matchings. A rich literature has developed since. See e.g., Roth and Sotomayor [1992] for an excellent survey; there is also a wealth of work that takes a more computational perspective [Manlove et al. 2010; Knuth 1997; Gusfield 1987; Bansal et al. 2010; Ashlagi et al. 2011; Hatfield and Kominers 2011; Manlove 1999, 2013; Gent et al. 2002; Irving

Permission to make digital or hardcopies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies show this notice on the first page or initial screen of a display along with the full citation. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credits permitted. To copy otherwise, to republish, to post on servers, to redistribute to lists, or to use any component of this work in other works requires prior specific permission and/or a fee. Permissions may be requested from Publications Dept., ACM, Inc., 2 Penn Plaza, Suite 701, New York, NY 10121-0701 USA, fax +1 (212) 869-0481, or permissions@acm.org.
EC'13, June 16–20, 2013, Philadelphia, USA. Copyright © 2013 ACM 978-1-4503-1962-1/13/06...$15.00

VISIT…

and Leather 1986; Irving et al. 1987; Manlove et al. 2002; Ashlagi et al. 2006; Bansal et al. 2010]. Overall, the study of matching has made contributions in both descriptive and prescriptive senses. Descriptively, computational models have enriched our understanding of existing social systems; e.g., Gale and Shapley's work helps us to understand marriage markets. Prescriptively, algorithmic tools proposed in the literature are practical enough to be used directly in applications, and indeed have transformed several matching markets. Perhaps most famously, the National Residency Matching Program (NRMP)—see, e.g., Roth [1996]—matches American medical students to internships at hospitals, using a method quite similar to the Gale–Shapley algorithm.

Much of the matching literature makes the key assumption that all market participants know their full (and often strict) preference orderings. This assumption is often reasonable, as demonstrated by the practical impact just described. However, as markets grow large it quickly becomes impractical for participants to assess their precise preference rankings. Returning to the NRMP, in practice students typically interview at only 10–15 hospitals out of about a thousand.The NRMP considers residents to have ranked as unacceptable all hospitals at which they did not interview, allowing the algorithm to proceed as though a full preference ordering had been declared. Prescribing the use of this system is clearly suboptimal: students who make the wrong decisions about where to interview can remain unmatched or be matched badly. There are likewise drawbacks from a descriptive perspective: such models shed little light on how matching works when participants are unsure about their preferences.

In this paper, we propose a novel model of two-sided matching in which participants are endowed with partially ordered preferences over candidates, but can refine these preferences through interviews. For example, these initial preferences might be ranked equivalence classes, with the property that candidates in higher-ranked equivalence classes are preferred to those in lower-ranked classes. Each participant's true (but unknown) preferences are a strict ordering over candidates, consistent with the known partial order and drawn from a known distribution. Participants can learn about their preferences through interviews. For a given participant, the first interview is uninformative; each subsequent interview reveals pairwise rankings among all previously interviewed candidates. Our goal is to design an adaptive, centralized interviewing policy which, given any partial information state, either selects an interview to perform or terminates with an optimal stable matching.

As interviews are costly, we would like our policy to minimize the number of interviews performed. We propose three criteria according to which the effectiveness of a policy can be assessed. First, a *(very weakly) dominant* policy performs a minimal number of interviews for all underlying true preference profiles consistent with the partial information. Second, a *Pareto optimal* policy may not be minimal for all underlying preference profiles, but is guaranteed not to be dominated by any other policy. Finally, an *optimal-in-expectation* policy minimizes the expected number of interviews with respect to the known distribution of strict preference orders. Note that the first two of these notions are prior free: i.e., such policies are independent of the preference distribution, and so are also applicable in more general settings where there is no common prior. Note also that, as dominant policies are Pareto optimal and optimal in expectation, the ideal is to compute a dominant policy. However, as we show, dominant policies may not exist.

Pareto optimal and optimal-in-expectation policies *are* guaranteed to exist; we show how to leverage a Markov Decision Process (MDP) encoding of our problem to find such policies in time polynomial in the number of possible preference orderings, which is exponentially faster than a naive algorithm, but still exponential in the size of the input. The key question we investigate in the rest of the paper is whether we can do better. We introduce the notion of an *optimality certificate*, which is a matching and a set of interviews whose outcomes can be used to prove that the matching is stable and optimal for a given side of the market. We show that in general, finding an optimality certificate involving the minimum number of interviews is NP-hard, providing evidence that finding an optimal-in-expectation and/or a Pareto dominant policy may be NP-hard. We next consider a restricted setting in which participants on one side of the market start out with the same equivalence classes structures, and in which pairs of agents must interview before they can match. We show that in this setting, minimum optimality certificates—and hence optimal policies—can be computed efficiently. Indeed, our policy is optimal in the strongest sense: it is dominant, meaning that it is also

prior free, Pareto optimal, and optimal in expectation with respect to any prior distribution that has full support. Lastly, we present preliminary investigations into extending this result to more general settings.

Related Work. We are aware of three threads of related work. The first augments the preference model to include indifference between candidates. For example, consider a school choice domain in which high school students are matched to public schools. Schools have some preferences over students based on geography and legacy considerations, but many students are equivalent according to these measures and hence it is reasonable to say that the school is indifferent between them. However, this literature (e.g., [Pathak and Sethuraman 2011; Abdulkadiroglu and Smez 2003]) is not intended to fully address the problem of incomplete preferences, and indeed it does not. Furthermore, even this simple extension to the standard model is quite tricky; for example, depending on how stability is defined, stable matchings are not always guaranteed to exist. There is an extensive body of work (see, e.g., Irving [1994], Irving and Manlove [2002], Irving et al. [2000], Manlove [1999], and Manlove [2002]) on studying three versions of stability that can be defined when indifference is permitted. Of these, the one most closely related to our work is that of super-stability. Loosely speaking, super-stable matchings are those that are stable no matter how the ties in the preference orderings are broken. In fact there is a relation between optimality certificates and super-stable matchings, which we discuss in Section 3. Polynomial time algorithms have been proposed for finding such matchings in various two-sided matching markets [Irving et al. 2000, 2003; Manlove 1999]).

A second thread of related work sets up the matching problem as a Bayesian game, assuming that participants have strict preferences drawn according to a commonly known probability distribution over preference orderings, and that all participants know their own preferences once the draw has taken place. Work in this vein typically seeks to design Bayes–Nash incentive compatible mechanisms or to describe the Bayes–Nash equilibria of standard mechanisms (see e.g. Roth and Sotomayor [1992]). One recent paper augments such Bayesian models with interviews, which are costly and chosen in a decentralized fashion [Lee and Schwarz 2009]. The authors investigate equilibrium interviewing behavior and observe that "clustered interviews" yield high social welfare. However, they do not insist that the final matching be stable with respect to agents' true preferences.

A third thread of work tries to derive properties of the market based on the available partial information. In recent work, Echenique et al. [2013] study settings in which agents' preferences are unobserved. They characterize matchings that are *rationalizable*—i.e., stable w.r.t. some underlying (unobserved) preference, as well as those that are rationalizable and optimal for one side of the market. An assumption crucial to most of their results is that the agents on each side of the market have different (unobserved) preference orderings, and it is known how many agents in the market have each possible preference ordering. Their work differs from ours in several ways; notably, we allow for preferences to be partially observed, and look for matchings that are stable and optimal for a given side of the market w.r.t. all possible underlying preference orderings (as opposed to only one). In a recent working paper, Haeringer and Iehle [2012] study matching markets in which the preferences of one side of the market are known, but the agents on the other side of the market only known whom they find acceptable. Haeringer and Iehle provide a technique for identifying pairs that are not matched in any stable matching for any underlying preference ordering, and hence for identifying unstable matchings.

2. OUR MODEL

Let $A = \{a_1, \ldots, a_n\}$ be a set of applicants and let $E = \{e_1, \ldots, e_m\}$ be a set of employers. We use the term *agents* when making statements that apply both to applicants and employers, and the term *candidates* to refer to agents on the side of the market opposite to that of an agent currently being considered. We assume that each employer can hire at most one applicant, and each applicant can be hired by at most one employer, hence focusing on one-to-one matching markets. Agents start out only partially aware of their preferences. Formally, each agent is initially aware of a strict

partial preference ordering over (a subset of) the candidates. We denote by p_{e_i} and p_{a_j} the strict partial preference ordering of e_i and a_j, respectively. We let $p_{E,A} = (p_{e_1}, \ldots, p_{e_m}, p_{a_1}, \ldots, p_{a_n})$ and call $p_{E,A}$ a partial preference ordering profile. For example, agents might start out by assigning candidates to equivalence classes, and having a strict preference ordering over these equivalence classes. This *equivalence class ordering* is a natural model for scenarios in which each agent knows that some candidates are her top-tier candidates, that others are her second-tier candidates, and so on. Figure 1 of Example 2.4 depicts such a setting, with each agent's partial preference ordering partitioning the candidates into strictly ranked equivalence classes.

Agents' true preferences are strict total orders: each applicant a has a strict preference ordering $\succ_a$ over $E \cup \{\varnothing\}$, and each employer e has a strict preference ordering $\succ_e$ over $A \cup \{\varnothing\}$. If an agent i prefers $\varnothing$ to candidate j, we say j is *unacceptable* to i; all other candidates are *acceptable* to i.[1] We let $\succ_{E,A} = (\succ_{e_1}, \succ_{e_2}, \ldots, \succ_{e_m}, \succ_{a_1}, \succ_{a_2}, \ldots, \succ_{a_n})$. The preference orderings $\succ_{E,A}$ are drawn according to a distribution whose support is consistent with the partial preference ordering profile. That is to say, for each agent i, total order $\succ$ is in the support of the distribution, and for every pair of candidates j and k: (i) if i prefers j to k according to the partial preference ordering profile then i strictly prefers j to k under $\succ$; and (ii) candidate j appears in i's partial preference ordering if and only if j is acceptable to i under $\succ$. Throughout the paper, we assume that the partial preference ordering profile $p_{E,A}$ is known and that the strict total order profile $\succ_{E,A}$ is unknown. Furthermore, most of our definitions and results do not assume knowledge of the distributions, and are thus *prior free*; i.e., the definitions and results apply regardless of the preference distribution.

2.1. Optimal stable matchings

The goal of our work is to construct a matching for agents that is stable with respect to the underlying preferences, and optimal for one side of the market. We now define these notions mathematically.

Definition 2.1 (*Matching*). A matching $\mu : A \cup E \mapsto A \cup E \cup \{\varnothing\}$ is an assignment of applicants to employers such that each applicant is assigned to at most one employer and vice versa. More formally, $\mu(a_j) = e_i$ if and only if $\mu(e_i) = a_j$, $\forall a_j \in A$ either $\exists e_i \in E, \mu(a_j) = e_i$ or $\mu(a_j) = \varnothing$ (the applicant is unmatched), and likewise $\forall e_i \in E$ either $\exists a_j \in A, \mu(e_i) = a_j$ or $\mu(e_i) = \varnothing$.

Definition 2.2 (*Blocking pair*). A pair (a_j, e_i) is a *blocking pair* in matching μ if a_j and e_i are not matched together in μ, $e_i \succ_{a_j} \mu(a_j)$, and $a_j \succ_{e_i} \mu(e_i)$.

Definition 2.3 (*Stable matching*). A matching μ is *stable* if it has no blocking pair and if no agent is matched to an unacceptable partner.

Example 2.4. Consider the setting with 2 employers and 2 applicants depicted in Figure 1. The column corresponding to each agent i gives that agent's strict partial preference ordering, which is in fact an equivalence class ordering, with the most preferred equivalence class at the top. In this example applicants have full knowledge of their preferences, while employers have no knowledge of their preferences. Table 2 gives all four possible strict preference orderings for the employers. In every case, matching μ_1, $\mu_1(e_1) = a_1$ and $\mu_1(e_2) = a_2$, is stable. Matching μ_2, $\mu_2(e_1) = a_2$ and $\mu_2(e_2) = a_1$, is also stable under (c). μ_2 is not stable in the other cases: (e_1, a_1) blocks μ_2 under (a) and (b), while (e_2, a_2) blocks μ_2 under (d).

Employer-optimal and applicant-optimal matchings are particularly interesting: these are the matchings most preferred by all employers (resp., applicants), as compared to all other stable matchings. When agents have strict preferences, as in our model, such matchings always exist and are unique [Gale and Shapley 1962]. These matchings can be used to build mechanisms in which truthtelling is a dominant strategy for the side of the market favored by the matching [Dubins and Freeman 1981; Roth 1982]. For example, if the employer-optimal matching is always picked, it is a dominant strategy for employers to reveal their preference orderings truthfully. In what follows

[1] We assume that agents have strict preferences over unacceptable candidates only to simplify notation.

$\mathbf{e_1}$	$\mathbf{e_2}$
a_1	a_1
a_2	a_2

$\mathbf{a_1}$	$\mathbf{a_2}$
e_1	e_2
e_2	e_1

Fig. 1. A setting with 2 employers and 2 applicants. Applicants have full knowledge of their preferences; employers don't.

$\mathbf{e_1}$	$\mathbf{e_2}$
a_1	a_1
a_2	a_2

(a)

$\mathbf{e_1}$	$\mathbf{e_2}$
a_1	a_2
a_2	a_1

(b)

$\mathbf{e_1}$	$\mathbf{e_2}$
a_2	a_1
a_1	a_2

(c)

$\mathbf{e_1}$	$\mathbf{e_2}$
a_2	a_2
a_1	a_1

(d)

Fig. 2. The four possible underlying preference profiles for the employers in the setting of Figure 1

we restrict our attention to the problem of finding employer-optimal matchings rather than arbitrary stable matchings.[2]

Definition 2.5 (*Employer-optimal matching*). A matching is employer-optimal if it is stable and weakly preferred by all employers to every other stable matching.

Example 2.6. Continuing Example 2.4, μ_1 is the employer-optimal matching under cases (a), (b), and (d), and μ_2 is the employer-optimal matching under case (c).

2.2. Policies

We have defined stability of a matching in terms of agents' true preferences (as distinct from Lee and Schwarz [2009]); thus, it will generally not be possible to find a stable matching based on agents' initial knowledge. To find and certify a stable matching, employers and applicants must gather additional information about each other through *interviews*. Each interview pairs a single applicant a with a single employer e; we denote such an interview $e{:}a$. After interview $e{:}a$, both e and a are able to strictly order the acceptable candidates interviewed so far.

Informally, a *policy* describes how to conduct interviews: it starts with input $I = (E, A, p_{E,A})$, performs a sequence of interviews, and then outputs a matching. Interviews can be selected based on the results of previous interviews and the whole procedure is centralized. More formally, a policy maps information states to subsequent interviews or to a matching that is stable and employer-optimal for the underlying preference profile. To reduce notation, from this point on we assume that the input $I = (E, A, p_{E,A})$ is given and fixed.

Definition 2.7 (*Information State*). The *information state* $\mathcal{I}_i$ of agent i after interviews with $\ell \geq 0$ candidates is a list of these ℓ candidates, ordered according to the underlying true preference profile. When $\ell = 0$, we let $\mathcal{I}_i = \varnothing$. The *global information state* after a sequence of interviews is $\mathcal{I} = \bigcup_{i \in E \cup A} \mathcal{I}_i$.

We say that information state $\mathcal{I}$ *refines* partial preference ordering profile $p_{E,A}$, which we write as $\mathcal{I} \lhd p_{E,A}$, if $\mathcal{I}$ is consistent with $p_{E,A}$: for all agents i and all candidates j and k, i preferring j to k under $p_{E,A}$ implies i also preferring j to k under $\mathcal{I}$. Given a fixed $p_{E,A}$ and information states $\mathcal{I}$ and $\mathcal{I}'$ that refine $p_{E,A}$, we say that $\mathcal{I}'$ refines $\mathcal{I}$ when it contains strictly more information: when for all agents i, all candidates who were interviewed in $\mathcal{I}_i$ were also interviewed in $\mathcal{I}'_i$, and for all candidates j and k who were interviewed in both $\mathcal{I}_i$ and $\mathcal{I}'_i$, i preferring j to k under $\mathcal{I}_i$ implies i preferring j to k under $\mathcal{I}'_i$. We say that a preference profile $\succ$ refines an information state $\mathcal{I}$, and write $\succ \lhd \mathcal{I}$, if $\succ$ is consistent with both $p_{E,A}$ and $\mathcal{I}$. We write $\succ \lhd p_{E,A}$ to denote that $\succ$ is consistent with $p_{E,A}$.

Definition 2.8 (*Policy*). A *policy* is a mapping from a global information state $\mathcal{I} \lhd p_{E,A}$ either to an interview to perform or to a matching. A policy is *sound* if it is guaranteed to return an employer-optimal matching, regardless of the true preference order $\succ \lhd p_{E,A}$.

Notice that a sound policy is allowed to return a matching in which a pair of agents e and a are matched even without e having interviewed a. In practice, employers may face the requirement of

[2] Of course, all of our technical results can be made to apply instead to applicant-optimal matchings by swapping every use of the terms "employer" and "applicant." Furthermore, all of our results except those in Section 5 also hold for the problem of finding *any* stable matching; however, we make the restriction throughout for consistency and clarity.

interviewing an applicant before offering her a job. We capture this additional requirement through the notion of a *diligent policy*.

Definition 2.9 (*Diligent policy*). A policy is *diligent* if it is sound and, furthermore, it maps from $\mathcal{I}$ to μ only if $e \in \mathcal{I}_a$ and $a \in \mathcal{I}_e$ for all $(e, a) \in \mu$.

Example 2.10. Continuing Example 2.6, a sound policy follows: given an information state $\mathcal{I}$, if e_1 has not interviewed a_1, schedule that interview; else if e_1 has not interviewed a_2, schedule that interview. Otherwise, if e_1 prefers a_2 to a_1: if e_2 has not interviewed a_1 schedule that interview; else if e_2 has not interviewed a_2, schedule that interview. These interviews are sufficient to distinguish between the underlying preference orderings that give rise to different employer-optimal matchings; hence, for the remaining information states, the policy can simply output a matching. If e_1 prefers a_1 to a_2, the true underlying ordering is (a) or (b), and so return μ_1. If e_1 prefers a_2 to a_1 and e_2 prefers a_2 to a_1, the true underlying ordering is (d), and so again return μ_1. Otherwise, the true underlying ordering is (c), and so return μ_2.

The above sound policy can be made into a diligent policy by the following modification: If e_1 prefers a_1 to a_2, have e_2 interview a_2 first and then return μ_1.

2.3. Minimizing the number of interviews

It is not hard to identify a sound policy: e.g., simply instruct all employers to interview all applicants and then run the Gale–Shapley algorithm to find the employer-optimal matching. However, this policy is likely to perform unnecessary interviews. We are motivated by the intuition that in reality, interviews are very costly; thus, we seek sound policies that minimize their number. However, there is a problem. Because policies select interview schedules dynamically, the number of necessary interviews may depend not only on the policy's input, but also on the true underlying preference orderings. What does it mean, then, to say that a sound policy performs the minimal number of interviews given an input I?

One straightforward answer is to say that we should minimize the expected number of interviews with respect to the prior distribution over preference orderings. Let $\theta(f, \succ, p_{E,A})$ denote the number of interviews that policy f performs when the true underlying preference profile is $\succ \lhd p_{E,A}$.

Definition 2.11 (*Optimal in expectation*). A policy f is *optimal in expectation* if it is sound and it minimizes the expected number of interviews performed, given the prior $\Pr$. That is, for all sound policies g, $\sum_{\succ \lhd p_{E,A}} \Pr(\succ) \cdot \theta(f, \succ, p_{E,A}) \leq \sum_{\succ \lhd p_{E,A}} \Pr(\succ) \cdot \theta(g, \succ, p_{E,A})$.

This objective has a major drawback: optimal-in-expectation policies are not robust with respect to changes in the setting, i.e., they are not prior-free. We would prefer a sound policy that performs no more interviews than any other sound policy, regardless of the underlying preference profile.

Definition 2.12 (*Very weak domination*). A policy f *very weakly dominates* another sound policy g if and only if f performs no more interviews than g for any underlying preference profile. That is, $\theta(f, \succ, p_{E,A}) \leq \theta(g, \succ, p_{E,A})$ for all $\succ \lhd p_{E,A}$.

This dominance notion is called "very weak" because two algorithms can very weakly dominate each other by always performing the same number of interviews for every $\succ$.

Definition 2.13 (*Very weakly dominant policy*). A policy f is *very weakly dominant* if it is sound and it very weakly dominates any other sound policy g.

We would like to find a very weakly dominant policy. Unfortunately, such policies do not always exist.

THEOREM 2.14. *There exist inputs for which no very weakly dominant policy exists.*

PROOF. Consider the setting given in Figure 1. To certify that μ_1 is employer-optimal for (a) or (b), we only need e_1 to interview both candidates a_1 and a_2, to distinguish (a) from (c). To

certify that μ_1 is employer-optimal for (b) or (d), we only need e_2 to interview both candidates, to distinguish (d) from (c). Thus any policy that instructs e_1 to interview first—e.g. the policy described in Example 2.10—is Pareto dominated in case (d), and any policy that instructs e_2 to interview first is Pareto dominated in case (a). □

Motivated by this impossibility result, we consider the weaker, but still prior-free, notion of Pareto optimality.

Definition 2.15 (*Pareto domination*). A policy f *Pareto dominates* another policy g if and only if both policies are sound, f very weakly dominates g and, furthermore, $\exists \succ \lhd p_{E,A}$ such that $\theta(f, \succ, p_{E,A}) < \theta(g, \succ, p_{E,A})$.

Definition 2.16 (*Pareto optimal policy*). A policy f is *Pareto optimal*[3] if there does not exist another policy g that Pareto dominates f.

For example, the policy described in Example 2.10 is Pareto optimal. Because of the strict inequality, Pareto domination is an asymmetric relation. Thus we can not have cycles in Pareto domination, and so

PROPOSITION 2.17. *A Pareto optimal policy always exists.*

Now we briefly survey relationships between our solution concepts. Very weak dominance is the strongest guarantee we can hope for: if we can find such a policy, it will also be Pareto optimal and optimal in expectation. Furthermore, every optimal-in-expectation policy is guaranteed to be Pareto optimal when the prior distribution has full support.

PROPOSITION 2.18. *If a policy f is optimal in expectation and* $\Pr$ *has full support, then f is Pareto optimal.*

PROOF. Assume for contradiction that f is not Pareto optimal. Thus there exists a policy g that Pareto dominates f. Therefore, $\theta(g, \succ, p_{E,A}) \leq \theta(f, \succ, p_{E,A})$ for all $\succ \lhd p_{E,A}$ and $\theta(g, \succ, p_{E,A}) < \theta(f, \succ, p_{E,A})$ for at least one $\succ \lhd pE, A$. Since $\Pr$ has full support, $\Pr(\succ) > 0$ for all $\succ \lhd p_{E,A}$. Therefore, $\sum_{\succ \lhd p_{E,A}} \Pr(\succ) \cdot \theta(f, \succ, p_{E,A}) > \sum_{\succ \lhd p_{E,A}} \Pr(\succ) \cdot \theta(g, \succ, p_{E,A})$, which implies that f does not minimize the expected number of interviews performed, a contradiction. □

All of the definitions in this section are stated in terms of sound policies. All of this section's definitions and results can be extended to diligent policies simply by replacing every occurrence of "sound" with "diligent". For example, f is a *very weakly dominant, diligent* policy if it is diligent and it very weakly dominates any other diligent policy g. Unfortunately, the nonexistence result of Theorem 2.14 can be extended to very weakly dominant, diligent policies.[4]

THEOREM 2.19. *There exist inputs for which no very weakly dominant, diligent policy exists.*

PROOF. Consider the same setting that we analyzed in the proof of Theorem 2.14, given in Figure 1. Note that at least one employer has to interview both applicants, or we can not distinguish between the four possible cases. Furthermore, all matched agents should interview each other, hence the other employer must interview at least one applicant. Therefore, any diligent policy performs at least three interviews before returning a matching.

To certify that $\mu_1 = \{(e_1, a_1), (e_2, a2)\}$ is employer-optimal for (a), we need e_1 to interview both candidates a_1 and a_2, to distinguish (a) from (c). Then, to have all matched agents interviewed each

[3]Pareto optimality is an extension of the standard concept of optimality to multiobjective settings. In game theory, the notion is traditionally used to reason about settings in which there is one objective function for each agent's utility function. We apply the notion differently: we do not have an objective function per agent, but rather an objective function (counting the number of interviews performed) per strict preference profile that refines the given partial information.

[4]Indeed, in what follows we will often observe that results that hold for sound policies also hold for diligent policies. However, important differences between the two kinds of policies emerge in Section 5.

other, we only need to additionally have e_2 to interview a_2. To certify that μ_1 is employer-optimal for (d), we need e_2 to interview both candidates, to distinguish (d) from (c). Then, to have all matched agents interviewed each other, we only need to additionally have e_1 interview a_1. Thus any diligent policy that instructs e_1 to interview first is Pareto dominated in case (d) and any diligent policy that instructs e_2 to interview first is Pareto dominated in case (a). □

3. FINDING OPTIMAL POLICIES

To compute an optimal-in-expectation (or Pareto optimal) policy, we can perform a brute-force search, considering every policy in turn. If S denotes the set of global information states, then the number of distinct policies is $\Theta((n^2)^{|S|})$—w.l.o.g., we assume that $n \geq m$. Thus, brute-force search requires time exponential in the number of information states, or $O((n^2)^{((n!)^n)})$. In this section we show how to do better.

We begin by identifying information states in which a sound policy can terminate and return a matching; i.e., in which enough information has been gathered that a policy can be sure of the employer-optimal matching.

Definition 3.1 (*Optimality certificate*). A triple $(\mathcal{I}, \mu, p_{E,A})$ is an *optimality certificate* if μ is employer-optimal w.r.t. all $\succ \lhd \mathcal{I}$.

In other words, an optimality certificate is an information state that admits a super-stable matching that is employer-optimal for all total orders consistent with the information state.

Define the size of an optimality certificate $(\mathcal{I}, \mu, p_{E,A})$ as the number of interviews performed in $\mathcal{I}$. A *minimal optimality certificate* is an optimality certificate that cannot be made smaller by dropping an interview.

Definition 3.2 (*Minimal optimality certificate*). A triple $(\mathcal{I}, \mu, p_{E,A})$ is a *minimal optimality certificate* if μ is employer-optimal w.r.t. all $\succ \lhd \mathcal{I}$, and if there is no smaller optimality certificate $(\mathcal{I}', \mu, p_{E,A})$ such that $\mathcal{I} \lhd \mathcal{I}'$.

Notice that a triple $(\mathcal{I}, \mu, p_{E,A})$ could be an optimality certificate even if a pair of agents, say e and a, are matched in μ but have not interviewed together according to $\mathcal{I}$. To certify the outputs of diligent policies, we also define *diligent optimality certificates*.

Definition 3.3 (*Diligent optimality certificate*). A triple $(\mathcal{I}, \mu, p_{E,A})$ is a *diligent optimality certificate* if μ is employer-optimal w.r.t. all $\succ \lhd \mathcal{I}$ and, furthermore, $e \in \mathcal{I}_a$ and $a \in \mathcal{I}_e$ for all $(e, a) \in \mu$.

Paralleling our earlier definitions, a *diligent minimal optimality certificate* is a diligent optimality certificate of minimal size.

THEOREM 3.4. *An optimal-in-expectation policy can be computed in time polynomial in* $|S|$.

PROOF. We leverage the planning paradigm of MDPs [see, e.g., Puterman 1994]. An MDP is a tuple $(S, \mathcal{A}, C, T, s_0, F)$, where S is a finite set of states; $\mathcal{A}$ is a finite set of actions; C is a cost function where $C(s, i, s')$ represents the cost of taking action i in state s and transitioning to state s'; T is a transition function where $T(s, i, s')$ denotes the probability that action i in state s leads to state s'; s_0 denotes the system's initial state; and F denotes a set of terminal states. Intuitively, our MDP encoding will work as follows. We start at an empty global information state and take actions corresponding to interviews, paying a cost of 1 for every interview performed. The effect of an action is to transition to a new information state that refines the previous state in the appropriate way, with new rankings being revealed according to conditional probabilities derived from the prior. We reach a terminal state when we have gathered enough information to stop conducting interviews. Formally, let S denote the set of global information states, $\mathcal{A}$ denote the set of all possible interviews, and $C(\mathcal{I}, e{:}a, \mathcal{I}') = 1$ for all $e{:}a$, $\mathcal{I}$ and $\mathcal{I}'$. Let $T(\mathcal{I}, e{:}a, \mathcal{I}') = \frac{\Pr(\mathcal{I}')}{\Pr(\mathcal{I})}$ if $\mathcal{I}'$ refines $\mathcal{I}$ and has exactly one more interview than $\mathcal{I}$, that interview being $e{:}a$, and $T(\mathcal{I}, e{:}a, \mathcal{I}') = 0$ otherwise, where

$\Pr(\mathcal{I}) = \sum_{\succ \lhd \mathcal{I}} \Pr(\succ)$. Let s_0 be the empty global information state: $\mathcal{I}$ where $\mathcal{I}_i = \varnothing$ for all i. Finally, we will define F as the set of information states such that $\mathcal{I} \in F$ if all preference profiles that refine $\mathcal{I}$ have the same employer-optimal matching, and there is no $\mathcal{I}'$, $\mathcal{I} \lhd \mathcal{I}'$ for which the same property holds.

We compute F as follows. We will iteratively refine a mapping h from information states to matchings, which will store all information states that we know to be safe stopping points for a sound policy. (In fact, h stores minimal optimality certificates.) Initially, let h map every information state to the null matching, in which no agent is matched. A sound policy can clearly stop when all interviews have been performed. Thus, for all information states $\mathcal{I}$ in which all interviews have been performed—including even interviews with unacceptable candidates—let $h(\mathcal{I}) = \mu_{\mathcal{I}}$, where $\mu_{\mathcal{I}}$ is the employer-optimal matching for the preference profile induced by $\mathcal{I}$. Initialize Q to be the set of all such information states, and initialize F to be the empty set. We repeat the following until Q is empty. Select an arbitrary information state $\mathcal{I}$ from Q such that $h(\mathcal{I})$ does not map to the null matching. Now, make a linear pass over Q, asking whether there exists a second information state $\mathcal{I}'$ such that $h(\mathcal{I}) = h(\mathcal{I}')$ and there exists a pair (e, a) where removing e from $\mathcal{I}_a$ and $\mathcal{I}'_a$ and removing a from $\mathcal{I}_e$ and $\mathcal{I}'_e$ yields the same global information state $\mathcal{I}^*$. If so, remove both $\mathcal{I}$ and $\mathcal{I}'$ from Q, add $\mathcal{I}^*$ to Q and let $h(\mathcal{I}^*) = h(\mathcal{I})$. If no such $\mathcal{I}'$ exists, remove $\mathcal{I}$ from Q and add it to F. The computation of F takes polynomial time in the number of states since initializing Q can be done in polynomial time, each pass over Q takes time $O(n \cdot m \cdot |S|)$, and at least one information state is removed from Q in each pass.

Our MDP is *finite horizon* because we always reach a terminal state within $n \cdot m$ actions (i.e., performing all interviews). Therefore, a standard result from the literature on MDPs applies: a policy that minimizes expected cost can be computed in time linear in the number of states via the backward induction algorithm (see, e.g., [Puterman and Patrick 2010]). This policy is sound by the construction of F, and hence is optimal in expectation. □

An optimal-in-expectation *diligent* policy can be computed in time polynomial in $|S|$. The proof is almost identical to the one given for Theorem 3.4 with the modifications that (1) h stores minimal diligent optimality certificates, and (2) we remove $\mathcal{I}$ and $\mathcal{I}'$ from Q and add $\mathcal{I}^*$ if, in addition to the constraints stated in the proof, the pair (e, a) does not belong to $h(\mathcal{I}) = \mu_{\mathcal{I}}$.

Overall, our MDP formulation yields an exponential time algorithm for identifying an optimal-in-expectation policy, because the size of S is $O(n!)^n$. While this is an exponential improvement over brute-force search, we would prefer a polynomial-time guarantee. (We do note, however, that one advantage of MDP formulations—in our domain and elsewhere—is that well-structured problem instances can be solved tractably even when the problem is intractable in the worst case.) Our MDP formulation made little use of the structure of the matching problem, so we have reason to hope that better worst-case performance might be possible. One hurdle is that explicitly stating a policy requires exponential space, and hence at least exponential time, as the number of information states is exponential. However, one could hope to do better simply by *executing* an optimal policy rather than stating it explicitly. The rest of this paper asks whether there are provably better ways of executing optimal policies.

4. FINDING MINIMUM OPTIMALITY CERTIFICATES

We now turn to the computational problem of finding an optimality certificate (recall Definition 3.1) that certifies the employer-optimal matching for a given preference profile in as few interviews as possible. (Contrast with Definition 3.2, which only required the local property that an optimality certificate could not be reduced by dropping an interview.) Such an optimality certificate would have to be identified by any very weakly dominant policy, and would likely be useful for optimal-in-expectation policies as well.

Definition 4.1 (*Minimum optimality certificate for* $\succ$). $(\mathcal{I}, \mu, p_{E,A})$ is *a (diligent) minimum optimality certificate for a preference profile* $\succ$ if μ is employer-optimal w.r.t. $\succ$, $\succ \lhd \mathcal{I}$, and if there does not exist a smaller (diligent) optimality certificate $(\mathcal{I}', \mu, p_{E,A})$ such that $\succ \lhd \mathcal{I}'$.

THEOREM 4.2. *A sound (or diligent) policy f is very weakly dominant (and diligent) if and only if it computes a (diligent) minimum optimality certificate for every preference profile $\succ \lhd p_{E,A}$.*

PROOF. First, consider the case of sound policies. To prove the first direction, assume for contradiction that f is not very weakly dominant. Then there must exist a policy g and a preference profile $\succ$ such that $\theta(g, \succ, p_{E,A}) < \theta(f, \succ, p_{E,A})$. Let μ be $\succ$'s employer-optimal matching and $\mathcal{I}$ be the information state after performing interviews that g performs under $\succ$. For g to be sound, $(\mathcal{I}, \mu, p_{E,A})$ must be an optimality certificate. Thus f does not compute a minimum optimality certificate for $\succ$, a contradiction. To prove the second direction, assume for contradiction that a minimum optimality certificate is not computed for some preference profile $\succ$. Let $(\mathcal{I}, \mu, p_{E,A})$ be a minimum optimality certificate for $\succ$. Take the set of interviews corresponding to $\mathcal{I}$, Z, and order them arbitrarily. Let the policy g be as follows: if no interview has taken place, schedule the first interview in Z; else if the first interview has taken place, schedule the second interview, and so on until all interviews in Z has taken place; if all the interviews in Z has taken place and the current global information state is $\mathcal{I}$, return μ, else perform all possible interviews and return the employer-optimal matching of the corresponding total order. Clearly g is sound and computes the minimum optimality certificate for $\succ$. Thus $\theta(g, \succ, p_{E,A}) < \theta(f, \succ, p_{E,A})$ and therefore g dominates f on $\succ$, a contradiction. Exactly the same argument suffices for diligent policies. □

Unfortunately, the problem of computing a minimum optimality certificate is NP-hard.

Definition 4.3 (*Optimality Certificate (OC) decision problem*). Given $(E, A, p_{E,A})$, a preference profile $\succ$ that refines $p_{E,A}$ and a bound K, decide whether there exists an optimality certificate $(\mathcal{I}, \mu, p_{E,A})$ of size at most K where μ is employer-optimal w.r.t. $\succ$.

THEOREM 4.4. *The optimality certificate decision problem is NP-hard, even if the partial preference ordering is an equivalence class ordering.*

PROOF. The proof is by reduction from the feedback arc set (FAS) problem [Karp 1972]. Let $G = (V, D)$ be a directed graph with no self-loops or multiple arcs. The feedback arc set problem is to decide whether there exists a set of arcs C of size $\leq K$ such that C contains at least one arc from every directed cycle in G; i.e., such that removing all arcs in C breaks all directed cycles of G. We construct an instance of the optimality certificate (OC) problem from (G, K) as follows. For each vertex i in G, create employers e_i and e'_i and applicants a_i and a'_i. For each employer e_i, create a single equivalence class of acceptable candidates, containing applicant a_i and every a_j where (i, j) is an arc in D. For each employer e'_i, create a single equivalence class of acceptable candidates containing applicants a_i and a'_i. For each applicant a_i, create two equivalence classes of acceptable candidates. Let the top class contain e_i and e'_i, and let the second class contain any e_j where (j, i) is an arc in D. For each applicant a'_i, create a single equivalence class of acceptable candidates containing only e'_i. Finally, let $\succ$ be any preference profile under which each e_i most prefers the corresponding a_i, each employer e'_i most prefers a'_i, each applicant a_i ranks e_i at top, and each applicant a'_i most prefers e'_i. Observe that matching μ where $\mu(e_i) = a_i$ and $\mu(e'_i) = a'_i$ is employer-optimal w.r.t. $\succ$. To complete the proof we show that G has a FAS of size at most K if and only if there exists an optimality certificate of size at most $K + 2n$ for $\succ$. The proof has two parts:

(1) Let $(\mathcal{I}, \mu, p_{E,A})$ be an optimality certificate of size K'. Let S be a set of arcs such that $(v_i, v_j) \in D$ iff $e_i{:}a_j \in \mathcal{I}$. We claim that S is of size at most $K' - 2n$ and removing all the arcs in S breaks all the cycles in G.
(2) Let S be a solution of size K to the feedback arc set problem. Let $\mathcal{I}$ be an information state consistent with $\succ$ under which each e_i, $1 \leq i \leq n$, has interviewed a_i and all applicants a_j such that $(v_i, v_j) \in S$, and furthermore, each employer e'_i, $\forall 1 \leq i \leq n$, has interviewed a_i. We claim that $(\mathcal{I}, \mu, p_{E,A})$ is an optimality certificate of size $K + 2n$ for $\succ$.

Proof of **(1)**: by contradiction. First note that $\mathcal{I}$ must include either both interviews $e_i{:}a_i$ and $e'_i{:}a_i$, or both interviews $e'_i{:}a_i$ and $e'_i{:}a'_i$, $\forall 1 \leq i \leq n$. As otherwise, an arbitrary total ordering

$\succ \lhd \mathcal{I}$ exists under which a_i and e'_i rank each other at the top and hence must be matched in the employer-optimal matching; contradicting that $(\mathcal{I}, \mu, p_{E,A})$ is an optimality certificate. Thus, there are at most $K' - 2n$ interviews in $\mathcal{I}$ that are of type (e_i, a_j). So S is of size at most $K' - 2n$.

Assume that removing S does not break all the cycles in G. Let C be an unbroken cycle. For each edge (v_i, v_j) in C it must be the case that $e_i{:}a_j \notin \mathcal{I}$. Let μ' be a matching in which $\mu'(e_i) = \mu(e'_i) = a'_i$ if e_i is not in C, and $\mu'(e_i) = \mu(e_j)$ if (e_i, e_j) is in C. Let $\succ'$ be a strict ordering under which each employer e'_i likes a'_i the best and each employer e_i not in C likes a_i the best (as in $\succ$), and each e_i in C likes a_j the best, where (e_i, e_j) is in C. Note $\succ'$ is consistent with $\mathcal{I}$. Furthermore, employers not in C have the same applicant on the top of their list as in $\succ$ (which should be consistent with the outcome of the interviews in $\mathcal{I}$), and for employers in C their true preferences between their partners in μ and μ' haven't been revealed yet (since no interview between such an employer and his match under μ' has taken place). Note that μ' is the employer-optimal matching for $\succ'$ and so μ is not the employer-optimal matching for $\succ'$. Thus $(\mathcal{I}, \mu, p_{E,A})$ is not an optimality certificate for $\succ$, a contradiction.

Proof of **(2)**: By contradiction. First note that it is easy to see, from the construction of $\mathcal{I}$, that $|\mathcal{I}| = K + 2n$. Assume that $(\mathcal{I}, \mu, p_{E,A})$ is not an optimality certificate for $\succ$. Thus, there exists a preference ordering $\succ'$ consistent with $\mathcal{I}$ for which μ is not the employer-optimal matching. We know that μ is a stable matching under any preference ordering consistent with $\mathcal{I}$ (it is the applicant-optimal matching indeed). Thus, it should be the case that there is some other stable matching μ' that the employers collectively prefer to μ (some are indifferent and some prefer μ'). Therefore, there has to be a set of employers $e_{c_1}, \ldots, e_{c_l}$ such that $\mu'(e_{c_i}) = \mu(e_{c_{i+1}})$ for $i < l$ and $\mu'(e_{c_l}) = \mu(e_{c_1})$. If not, then there exists an unmatched employer and unmatched applicant who both prefer to be matched together than stay unmatched, thus μ' is not stable. Note that e_{c_i} must prefer his match in μ' to his match in μ, a_{c_i}, under $\succ'$. Thus e_{c_i} has $\mu(e_{c_{i+1}})$ in his one and only equivalence class and therefore $(v_{c_i}, v_{c_{i+1}}) \in D$. Furthermore $(v_{c_l}, v_{c_1}) \in D$. Thus, $e_{c_1}, \ldots, e_{c_l}$ form a cycle in G, a contradiction. □

We can analogously define the diligent optimality certificate (DOC) decision problem. A corresponding hardness claim also holds for DOC. This proof is quite similar to the proof just given; we omit it to save space.

An optimal-in-expectation or a Pareto optimal policy need not compute optimality certificates for every input. Nevertheless, we consider the hardness of computing optimality certificates to be discouraging evidence about the tractability of the problem of identifying such policies.

5. IDENTICAL EQUIVALENCE CLASS ORDERINGS ON ONE SIDE OF THE MARKET

In many two-sided matching markets, there is some degree of positive correlation between agents' preferences. Taken to the extreme, this means that agents on at least one side of the market may have common *a priori* information, though perhaps different underlying preferences. In this section we focus on such a setting, and furthermore consider the special case of partial preference orderings that we introduced earlier in Section 2 (as "equivalence class orderings").

Specifically, we now assume that each agent has an equivalence relation over (a subset of) the candidates and a strict ordering over the equivalence classes. We denote the equivalence classes of employer e_i and applicant a_j that are ranked ℓth in this strict ranking by $c_{e_i,\ell}$ and $c_{a_j,\ell}$, respectively. We let $c_{E,A}$ denote the equivalence class profile for all employers and applicants. We assume that the applicants are endowed with identical equivalence class orderings. Thus, formally we assume that $c_{a_i,k} = c_{a_j,k}$ for all pairs of applicants a_i and a_j and all k; we allow applicants to have different distributions and different underlying preference orderings, i.e., $\succ_{a_i} \neq \succ_{a_j}$. We do not restrict employers' preferences in any way. Irving et al. [2008] study a similar setting in which the preference orderings of one or both sides of the market are derived from a common ranking of all the candidates into strictly ranked equivalence classes. They refer to this common ranking as a *master list*. They assume that each agent's preference ordering is then derived from the common ranking by eliminating his or her unacceptable partners, thus preserving the indifferences if they exist in the master list. They

were motivated in this work by how the applicants were ranked in the hospitals–residents matching market in England in 2005–2006. Their setting differs from ours in the sense that we allow, and even require, the agents to break ties.

We will obtain positive results in this restricted setting under the additional restriction to diligent policies. (We demonstrate the necessity of our restriction to diligent policies at the end of the section.) Even with all these restrictions, observe that a tabular representation of an optimal-in-expectation, diligent policy requires exponential space, since there are an exponential number of information states. As discussed earlier, we get around this problem by supplying an algorithm that *executes* such a policy rather than specifying it explicitly. Specifically, we present a polynomial-time algorithm that executes a very weakly dominant, diligent policy, and hence both an optimal-in-expectation and a Pareto optimal, diligent policy.

Our algorithm for identifying a very weakly dominant, diligent policy works like Gale–Shapley except that it is interleaved with another procedure that instructs agents to interview. It is described formally as Algorithm 1; an informal description follows. We say that employer e ranks applicant a in class ℓ if a is in the ℓth equivalence class of e. Similarly, we say that the applicants rank e in class ℓ if e is in the ℓth equivalence class of the applicants. Algorithm 1 alternates between two main stages, an interview stage and a tentative matching stage.

— In the interview stage, the algorithm chooses an unmatched employer e that satisfies two properties. First, e has achievable applicants that he has not yet interviewed. An achievable applicant is one that is acceptable to e, and has not yet received better offers. Second, if applicants rank e in class ℓ, then they rank any other employer e' that satisfies the first property in a class below or equal to ℓ. Employer e then interviews all achievable applicants that he has not yet interviewed in his equivalence class that is highest-ranked among those containing his achievable applicants.
— In the tentative matching stage, unmatched employers propose to their most preferred achievable, interviewed applicant, if any. Each applicant a tentatively accepts her best proposal, say e. When there is no unmatched employer with an achievable interviewed applicant, the tentative matching stage ends.

If there are still unmatched employers with achievable applicants that they have not yet interviewed, the algorithm returns to the interview stage. Otherwise, the algorithm halts and returns the final tentative matching. The next example shows how the algorithm works.

Example 5.1. Consider the setting with 3 employers and 3 applicants depicted in Figure 3. Notice that the applicants are endowed with identical equivalence classes. Assume that the true strict preference ordering of the agents is $\succ_1$, depicted in Figure 4. Running Algorithm 1 on this setting returns the employer-optimal stable matching $\mu_1 = \{(e_1, a_1), (e_2, a_3), (e_3, a_2)\}$ after 3 runs of the main loop and 5 interviews. Assume, for the sake of this example, that in the interview stage, Algorithm 1 chooses the lowest indexed employer that satisfies the two required properties. Then, the execution of Algorithm 1 on $\succ_1$ will be as follows.

(1) e_1 is instructed to interview both a_1 and a_2. e_1 proposes to his most favorite candidate a_1 and is matched to her. a_1 is no longer achievable for e_3 and so is removed from his list.
(2) e_2 is instructed to interview both a_1 and a_3. e_2 proposes to his most favorite candidate a_3 and is matched to her. a_3 is no longer achievable to e_3 and so is removed from his list.
(3) e_3 is instructed to interview a_2, the only applicant still achievable for him. e_3 proposes to a_2 and is matched to her.

If the true ordering was $\succ_2$, depicted in Figure 5, then running Algorithm 1 would return the employer-optimal stable matching $\mu_2 = \{(e_1, a_2), (e_2, a_1), (e_3, a_3)\}$ after 3 runs of the main loop and 5 interviews.

(1) e_1 is instructed to interview both a_1 and a_2. e_1 proposes to his most favorite candidate a_1 and is matched to her. a_1 is no longer achievable to e_3 and so is removed from his list.

Algorithm 1: Lazy Gale-Shapley

Input: E, A, $c_{E,A}$
Output: Matching μ
Initialize
  $\ell = 1$; // Current equivalence class number of applicants
  $\mu(e_i) \leftarrow \emptyset, \mu(a_j) \leftarrow \emptyset, \forall e_i \in E, \forall a_j \in A$;
  $\chi_{e_i} \leftarrow \emptyset, \forall e_i \in E$; // e_i's set of achievable interviewed applicants
repeat
  Interview Stage
  foreach $e_i \in E$ **do**
    if $\mu(e_i) = \emptyset$ **then**
      $P_{e_i} \leftarrow$ The set of achievable applicants in the highest-ranked equivalence class of e_i among those with his achievable applicants, whom he has not interviewed yet ;
  $E^{u,\ell} \leftarrow$ the set of unmatched employers e_d in equivalence class ℓ of the applicants with nonempty P_{e_d} ;
  while $E^{u,\ell} = \emptyset$ **do**
    $\ell \leftarrow \ell + 1$;
    $E^{u,\ell} \leftarrow$ the set of unmatched employers e_d in equivalence class ℓ of the applicants with nonempty P_{e_d} ;
  $e_i \leftarrow$ an arbitrary employer in $E^{u,\ell}$;
  foreach $a_j \in P_{e_i}$ **do**
    e_i interviews a_j ;
  $\chi_{e_i} = P_{e_i}$;
  $P_{e_i} \leftarrow \emptyset$;
  Tentative Matching Stage
  repeat
    foreach $e_i \in E$ **do**
      if $\mu(e_i) = \emptyset$ *and* $\chi_{e_i} \neq \emptyset$ **then**
        e_i offers a position to the achievable applicant it prefers the most; she is removed from χ_{e_i} ;
    Each applicant who has received one or more job offers tentatively accepts the offer from the employer she most prefers and rejects the rest; matching μ is updated accordingly ;
    Each tentatively matched applicant a_j is no more achievable to those employers that are in the equivalence classes ranked lower than the one $\mu(a_j)$ belongs to; she is removed from the lists of such employers; χ_{e_i} is updated accordingly for all $e_i \in E$;
  until *there is no unmatched employer* e_i *with* $\chi_{e_i} \neq \emptyset$;
until *each employer is either tentatively matched or has no achievable applicants*;
return μ ;

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{e_3}$
a_1	a_1	a_1
a_2	a_3	
a_3	a_2	a_2
		a_3

p_E

$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{a_3}$
e_1	e_1	e_1
e_2	e_2	e_2
e_3	e_3	e_3

p_A

Fig. 3. A setting with 3 employers and 3 applicants. Applicants are endowed with identical equivalence classes.

(2) e_2 is instructed to interview both a_1 and a_3. e_2 proposes to his most favorite candidate, a_1. a_1 accepts e_2's proposal and rejects e_1 whom she likes less than e_2. e_2 is matched to a_1 and a_1 is removed from e_1's list. a_1 proposes to a_2 who is now at the top of his list, and is matched to her. a_2 is no longer achievable to e_3 and so is removed from his list.
(3) e_3 is instructed to interview a_3, the only applicant still achievable to him. e_3 proposes to a_3 and is matched to her.

Note that $\succ_2$ is almost identical to $\succ_1$, depicted in Figure 4, except that e_2's top choices are reversed.

Our main claim in this section is that Algorithm 1 executes a very weakly dominant, diligent policy. Key to the proof is to show that, regardless of the underlying preference profile, Algorithm 1 always identifies a diligent minimum optimality certificate.

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{e_3}$
a_1	a_3	a_1
a_2	a_1	a_2
a_3	a_2	a_3

$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{a_3}$
e_2	e_1	e_1
e_1	e_2	e_2
e_3	e_3	e_3

Fig. 4. Agents' preferences under total order $\succ_1$ that refines the setting depicted in Figure 3.

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{e_3}$
a_1	a_1	a_1
a_2	a_3	a_2
a_3	a_2	a_3

$\mathbf{a_1}$	$\mathbf{a_2}$	$\mathbf{a_3}$
e_2	e_1	e_1
e_1	e_2	e_2
e_3	e_3	e_3

Fig. 5. Agents' preferences under total order $\succ_2$ that refines the setting depicted in Figure 3.

THEOREM 5.2. *Algorithm 1 executes a very weakly dominant, diligent policy in time polynomial in the input size.*

PROOF. The proof proceeds in four steps.

Step 1: *Algorithm 1 terminates in time polynomial in the input size.* Note that Algorithm 1 halts once each employer is matched or has no more achievable applicants to propose to or interview. Hence, the algorithm is guaranteed to perform at least at least one interview in the interview stage of each iteration. Furthermore, in each invocation of the tentative matching stage, at least one employer proposes to an applicant and is then either rejected by that applicant or is tentatively matched. There are $n \cdot m$ possible interviews and hence $n \cdot m$ possible proposals. Thus the algorithm halts in polynomial time.

Step 2: *Algorithm 1 executes a diligent policy.* Recall that the employer-proposal variant of the Gale–Shapley algorithm yields the employer-optimal stable matching. The algorithm is robust w.r.t. varying the order in which the employers propose, as long as no tentatively matched employer proposes, and employers always propose to their top choice among achievable candidates. In Algorithm 1, no employer interviews an applicant unless all the applicants he ranks in higher equivalence classes are unachievable, no employer proposes when he is tentatively matched, and unmatched employers always propose to the most-preferred achievable, interviewed applicant. The algorithm only halts when each employer is matched or has no more achievable applicants to propose to or interview. Thus Algorithm 1 returns the same matching that Gale–Shapley would have returned for any preference profile that refines the global information state that holds at the moment Algorithm 1 terminates. Because Gale–Shapley is sound, Algorithm 1 is sound. Since no employer proposes to an applicant he has not interviewed, Algorithm 1 is also diligent.

Step 3: *Algorithm 1 only ever instructs an employer in applicants' equivalence class ℓ to interview applicants, or to make offers, when all employers in applicants' higher-ranked equivalence classes are matched to their employer-optimal matches or have been rejected by all applicants they find acceptable.* By the interview stage, if e in applicants' equivalence class ℓ is chosen to perform interviews, all employers in higher-ranked equivalence classes must either be tentatively matched or have no more achievable applicants. If none of the matched employers makes a new offer, their tentative matches become final, and so must be, by Step 2, their employer-optimal matches. If any of the matched employers makes a new offer, it can only be because he has been rejected by his current tentative match. Let e_* be the first employer in applicants' equivalence class $\ell - 1$ or higher that gets rejected by his current match in a round in which employers in applicants' equivalence class ranked ℓ or lower are interviewing. Then e_*'s tentative match must have received an offer from a more-preferred employer, e'_*, in applicants' equivalence class $\ell - 1$ or higher. This is only possible if e'_* was rejected by his match in an earlier round, contradicting our definition of e_*.

Step 4: *Algorithm 1 computes a diligent minimum optimality certificate $(\mathcal{I}, \mu, c_{E,A})$ for any preference profile $\succ$ such that $\succ \lhd c_{E,A}$.* Given that Algorithm 1 is diligent, it is sufficient to show that all interviews performed by Algorithm 1 on preference profile $\succ$ are in $\mathcal{I}$. We show this in two steps. For a given employer e that applicants rank in class ℓ, let Ω be the set of all applicants in equivalence classes of e that are above or equal to the equivalence class containing $\mu(e)$, unless those applicants are matched by μ to employers that applicants rank in a class above ℓ. We first show that $\mathcal{I}$ includes all interviews in Ω. We then show that Algorithm 1 performs only interviews in Ω.

— *$\mathcal{I}$ includes all interviews in Ω.* By our requirement that matched agents must interview each other, any diligent policy must require e to interview $\mu(e)$. Furthermore, for each a, $a \neq \mu(e)$, who is ranked by e in the same equivalence class as $\mu(e)$ or higher, and who is not matched to an employer she ranks in a higher equivalence class than ℓ, there exists a preference ordering $\succ'$ that is the same as $\succ$ except that e and a promote each other to the highest possible position in their respective rankings. Under $\succ'$, (e, a) blocks μ; thus μ is not employer-optimal under $\succ'$. However, unless e interviews a, preference profile $\succ'$ refines $\mathcal{I}$ and thus $(\mathcal{I}, \mu, c_{E,A})$ is not an optimality certificate.
— *Algorithm 1 performs only interviews in Ω.* We need to show that e does *not* interview applicant a if applicants rank $\mu(a)$ in a class above ℓ (recall that ℓ is the applicants' class containing e), and also if e ranks a in a class below the class containing $\mu(e)$. By Step 3, any a who ranks $\mu(a)$ in an equivalence class ranked higher than ℓ must be matched to him when e is chosen and thus is not achievable to e and so is not interviewed by e. If e ranks a in a class below $\mu(e)$, then e also does not interview a in Algorithm 1 because e does not interview applicants in an equivalence class unless he is rejected by all applicants in higher ranked equivalence classes. Since $\mu(e)$ does not reject e, we can conclude that e does not interview a.

Finally, the desired result follows by Theorem 4.2. □

Recall that to get a polynomial-time result, we not only restricted our partial information setting, but also turned to diligent policies as opposed to sound policies. We have already seen, in Theorem 2.19, that without any restriction on the given partial preference ordering, we can not guarantee the existence of a very weakly dominant, diligent policy. The next theorem shows that even in our restricted setting, we can not guarantee the existence of a sound policy that very weakly dominates every other sound policy.

THEOREM 5.3. *There exist equivalence class orderings in which applicants are endowed with identical equivalence classes, for which no very weakly dominant policy exists.*

PROOF. Consider a setting with 2 employers and 2 applicants where all agents initially rank both candidates in the same equivalence class. Let $\mathcal{I}_1$, $\mathcal{I}_2$, $\mathcal{I}_3$, and $\mathcal{I}_4$ be four information states that can be reached after a sequence of three interviews with the properties that:

— under $\mathcal{I}_1$, e_1 and a_1 prefer each other the most,
— under $\mathcal{I}_2$, e_2 and a_2 prefer each other the most,
— under $\mathcal{I}_3$, e_2 and a_1 prefer each other the most, and
— under $\mathcal{I}_4$, e_1 and a_2 prefer each other the most.

The reader can easily verify that these four information states are indeed all reachable after three interviews in the setting described. It is also easy to verify that $\mu_1 = \{(e_1, a_1), (e_2, a_2)\}$ is the employer-optimal matching for any preference ordering that refines $\mathcal{I}_1$ and/or $\mathcal{I}_2$, and $\mu_2 = \{(e_1, a_2), (e_2, a_1)\}$ is the employer-optimal matching for any preference ordering that refines $\mathcal{I}_3$ and/or $\mathcal{I}_4$.

Let $\succ_i \lhd \mathcal{I}_i$, $1 \leq i \leq 4$, be the four strict preference orderings depicted in Figure 6. The minimum optimality certificate for $\succ_1$ is $(\mathcal{I}_1, \mu_1, p_{E,A})$. To see this, note that to certify μ_1 as the employer-optimal matching for $\succ_1$, e_1 needs to have interviewed both applicants, to distinguish between $\succ_1$ and $\succ' \lhd \mathcal{I}_4$, $e_1 \succ'_{a_1} e_2$ and $a_1 \succ'_{e_2} a_2$. In addition, e_2 has to interview a_1 in order to distinguish between $\succ_1$ and $\succ'' \lhd \mathcal{I}_3$, where $\succ$ and $\succ''$ are the same except that a_1's preferences are reversed. Similarly, we can prove that the minimum optimality certificate for $\succ_2$ is $(\mathcal{I}_2, \mu_1, p_{E,A})$, for $\succ_3$ is $(\mathcal{I}_3, \mu_2, p_{E,A})$, and for $\succ_4$ is $(\mathcal{I}_4, \mu_2, p_{E,A})$. Notice that each of these minimum optimality certificates are reached by three interviews. Any policy that performs the interview e_1:a_1 (respectively e_2:a_2, e_1:a_2, e_2:a_1) is Pareto dominated under $\succ_2$ (respectively under $\succ_1$, $\succ_3$, $\succ_4$). Hence, no very weakly dominant policy exists. □

The reader may wonder whether we can recover our positive result if we restrict the setting such that the employers, as opposed to applicants, are endowed by identical equivalence classes. (Observe that the restriction is not symmetric because in both cases we are concerned with identifying

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{a_1}$	$\mathbf{a_2}$
a_1	a_1	e_1	e_1
a_2	a_2	e_2	e_2

$\succ_1$

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{a_1}$	$\mathbf{a_2}$
a_2	a_2	e_2	e_2
a_1	a_1	e_1	e_1

$\succ_2$

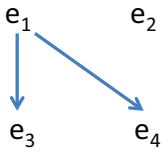
$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{a_1}$	$\mathbf{a_2}$
a_1	a_1	e_2	e_2
a_2	a_2	e_1	e_1

$\succ_3$

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{a_1}$	$\mathbf{a_2}$
a_2	a_2	e_1	e_1
a_1	a_1	e_2	e_2

$\succ_4$

Fig. 6. Four of the possible strict preference orderings that refine a setting with 2 employers and 2 applicants with empty initial information.



Applicants' partial preference orderings

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{e_3}$	$\mathbf{e_4}$
a_1	a_2	a_2	a_2
a_2	a_1	a_1	a_1

Employers' partial preference orderings

Fig. 7. A setting with 4 employers and 2 applicants. Employers have full knowledge of their preferences as shown in the table on right. Applicants' partial preference ordering is as depicted on the left: they both prefer e_1 to e_3 and e_4.

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{e_3}$	$\mathbf{e_4}$
a_1	a_2	a_2	a_2
a_2	a_1	a_1	a_1

$\mathbf{a_1}$	$\mathbf{a_2}$
e_1	e_2
e_2	e_1
e_3	e_3
e_4	e_4

Fig. 8. Agents' preferences under total order $\succ$.

$\mathbf{e_1}$	$\mathbf{e_2}$	$\mathbf{e_3}$	$\mathbf{e_4}$
a_1	a_2	a_2	a_2
a_2	a_1	a_1	a_1

$\mathbf{a_1}$	$\mathbf{a_2}$
e_1	e_1
e_2	e_2
e_3	e_3
e_4	e_4

Fig. 9. Agents' preferences under total order $\succ'$.

employer-optimal matchings.) It turns out that we can not; instead, this new restricted setting leads us back to the nonexistence result of Theorem 2.19.

THEOREM 5.4. *There exist equivalence class orderings in which employers are endowed with identical equivalence classes, for which no very weakly dominant, diligent policy exists.*

The proof follows immediately from the proof of Theorem 2.19. Lastly, we show that our restriction to equivalence class orderings was also important for getting us around the nonexistence of very weakly dominant, diligent policies.

THEOREM 5.5. *There exist inputs for which no very weakly dominant diligent policy exists, even when applicants are endowed with identical, strict partial preference orderings.*

PROOF SKETCH. Consider a setting with 4 employers and 2 applicants where the agents' partial preference orderings are as depicted in Figure 7. That is, the employers initially have full knowledge of their preferences and the applicants' identical partial preference orderings reveal that they both prefer e_1 to e_3 and e_4. Let the agents' preferences under total orders $\succ$ and $\succ'$ be as depicted in Figure 8 and Figure 9. Note that $\succ$ and $\succ'$ are the same except that a_2's top two choices are interchanged.

Matching μ, $\mu(e_1) = a_1$ and $\mu(e_2) = a_2$, is the employer-optimal matching under both $\succ$ and $\succ'$. The minimum diligent optimality certificate for $\succ$, $(\mathcal{I}, \mu)$, consists of the interviews $Z = \{e_1{:}a_1, e_2{:}a_2, e_1{:}a_2\}$, whereas the minimum diligent optimality certificate for $\succ'$, $(\mathcal{I}', \mu)$, consists of the interviews $Z' = \{e_1{:}a_1, e_2{:}a_2, e_3{:}a_2, e_4{:}a_2\}$. Note that Z and Z' only share the two interviews $e_1{:}a_1$ and $e_2{:}a_2$, and that both $\succ$ and $\succ'$ refine the information state resulting from these two interviews. Thus, a policy that computes a minimum diligent optimality certificate for $\succ$ cannot compute one for $\succ'$ (and vice versa): a policy that performs $e_1{:}a_2$ is Pareto dominated under $\succ'$, and a policy that performs $e_3{:}a_2$ and/or $e_4{:}a_2$ is Pareto dominated under $\succ$. Therefore, no very weakly dominant diligent policy exists for the given setting. □

6. CONCLUSIONS

We have introduced a model of two-sided matching markets in which agents begin with partially ordered preference information, and can refine these preferences through interviews. We defined three optimization criteria to capture the idea of minimizing the number of interviews required to find a stable matching that is optimal for a given side of the market. We showed that among these criteria, very weakly dominant policies do not always exist, and that in general settings it is NP-hard to find such a policy if it does exist. In contrast, optimal-in-expectation and Pareto-optimal policies do always exist; they can be found in exponential time via an MDP encoding. We can do better in the setting where one side of the market is endowed with identical equivalence class orderings; here, we can leverage the notion of minimum optimality certificates to execute a very weakly dominant, diligent policy in polynomial time.

Our paper raises many questions. A particularly important open problem is the hardness of finding optimal-in-expectation policies (or Pareto optimal policies) in general settings. One could also consider approximating these objectives. It would be interesting to identify settings under which we can bound the number of interviews our proposed algorithm performs (e.g., w.r.t. the number of equivalence classes and the number of agents in each equivalence class). In this vein, it would be particularly useful to understand when we only need a linear number of interviews, as this mimics the setting in the NRMP where applicants can only list a constant number of hospitals. Another interesting direction is to forgo the employer-optimality of the matching, and look for approximately employer-optimal matchings, in exchange for fewer interviews. Conversely, it is also interesting to investigate whether one can guarantee the employer-optimal matching by performing at most boundedly more than the minimum number of interviews required by an optimal policy. It is also important to investigate how and when our results could be extended to the many-to-one matching markets. This would of course depend on the employers' preferences over groups of applicants, as well as the differences in the number of applicants each employer intends to hire. For example, if employers have different number of spots available, it may be beneficial to have employers with larger numbers of spots interview first. Finally, it would be worthwhile to investigate the computational impact of committing to decentralized policies.

ACKNOWLEDGMENTS

We thank Scott Kominers for helpful conversations during the research that led to this paper, and anonymous reviewers for offering valuable feedback.

REFERENCES

ABDULKADIROGLU, A. AND SMEZ, T. 2003. School choice: A mechanism design approach. Discussion papers, Columbia University, Department of Economics.

ASHLAGI, I., BRAVERMAN, M., AND HASSIDIM, A. 2011. Matching with couples revisited. In *ACM-EC*. 335–336.

ASHLAGI, I., MONDERER, D., AND TENNENHOLTZ, M. 2006. Resource selection games with unknown number of players. In *AAMAS*. 819–825.

BANSAL, N., GUPTA, A., LI, J., MESTRE, J., NAGARAJAN, V., AND RUDRA, A. 2010. When lp is the cure for your matching woes: Improved bounds for stochastic matchings. *Algorithms–ESA 2010*, 218–229.

DUBINS, L. AND FREEMAN, D. 1981. Machiavelli and the Gale–Shapley algorithm. *American Mathematical Monthly 88,* 7, 485–494.

ECHENIQUE, F., LEE, S., SHUM, M., AND YENMEZ, M. B. 2013. The revealed preference theory of stable and extremal stable matchings. *Econometrica 81,* 1, 153–171.

GALE, D. AND SHAPLEY, L. 1962. College admission and the stability of marriage. *The American Mathematical Monthly 69,* 1, 9–15.

GENT, I. P., PROSSER, P., SMITH, B., AND WALSH, T. 2002. SAT encodings of the stable marriage problem with ties and incomplete lists. In *SAT 2002*. 133–140.

GUSFIELD, D. 1987. Three fast algorithms for four problems in stable marriage. *SIAM Journal of Computation 16,* 1, 111–128.
HAERINGER, G. AND IEHLE, V. 2012. Two-sided matching with one-sided information.
HATFIELD, J. AND KOMINERS, S. 2011. Multilateral matching. In *ACM-EC*. 337–338.
IRVING, R. W. 1994. Stable marriage and indifference. *Discrete Applied Mathematics 48,* 3, 261 – 272.
IRVING, R. W. AND LEATHER, P. 1986. The complexity of counting stable marriages. *SIAM J. of Computation 15*, 655–667.
IRVING, R. W., LEATHER, P., AND GUSFIELD, D. 1987. An efficient algorithm for the "optimal" stable marriage. *J. ACM 34*, 532–543.
IRVING, R. W. AND MANLOVE, D. F. 2002. The stable roommates problem with ties. *J. Algorithms 43,* 1, 85–105.
IRVING, R. W., MANLOVE, D. F., AND SCOTT, S. 2000. The hospitals/residents problem with ties. In *SWAT*. 259–271.
IRVING, R. W., MANLOVE, D. F., AND SCOTT, S. 2003. Strong stability in the hospitals/residents problem. In *Proceedings of STACS 2003: the 20th International Symposium on Theoretical Aspects of Computer Science*, H. Alt and M. Habib, Eds. Lecture Notes in Computer Science Series, vol. 2607. Springer-Verlag GmbH, 439–450.
IRVING, R. W., MANLOVE, D. F., AND SCOTT, S. 2008. The stable marriage problem with master preference lists. *Discrete Applied Mathematics 156,* 15, 2959–2977.
KARP, R. M. 1972. Reducibility among combinatorial problems. In *Complexity of Computer Computations*. 85–103.
KNUTH, D. 1997. *Stable marriage and its relation to other combinatorial problems: an introduction to the mathematical analysis of algorithms*. CRM proceedings & lecture notes. American Mathematical Society.
LEE, R. AND SCHWARZ, M. 2009. Interviewing in two-sided matching markets. *NBER Working Paper No. 14922*.
MANLOVE, D. F. 1999. Stable marriage with ties and unacceptable partners. Tech. rep., University of Glasgow, Department of Computing Science.
MANLOVE, D. F. 2002. The structure of stable marriage with indifference. *Discrete Applied Mathematics 122,* 13, 167 – 181.
MANLOVE, D. F. 2013. *Algorithmics of Matching Under Preferences*. World Scientific.
MANLOVE, D. F., IRVING, R. W., AND IWAMA, K. 2010. Special issue on matching under preferences. *Algorithmica, 58,* 1.
MANLOVE, D. F., IRVING, R. W., IWAMA, K., MIYAZAKI, S., AND MORITA, Y. 2002. Hard variants of stable marriage. *Theoretical Computer Science 276,* 1-2, 261 – 279.
PATHAK, P. AND SETHURAMAN, J. 2011. Lotteries in student assignment: An equivalence result. *Theoretical Economics 6,* 1, 1–17.
PUTERMAN, M. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley and Sons, New York.
PUTERMAN, M. AND PATRICK, J. 2010. *Encyclopedia of Machine Learning*. Chapter Dynamic programming.
ROTH, A. 1982. The economics of matching: stability and incentives. *Mathematics of Operations Research 7,* 4, 617–628.
ROTH, A. 1996. The national residency matching program as a labor market. *Journal of the American Medical Association 275,* 13, 1054–1056.
ROTH, A. AND SOTOMAYOR, M. 1992. *Two-Sided Matching: A Study in Game-Theoretic Modeling and Analysis*. Cambridge University Press.